

Explicit Implicit Function Theorem for All Fields

YINING HU

CNRS, Institut de Mathématiques de Jussieu-PRG

Université Pierre et Marie Curie, Case 247

4 Place Jussieu

F-75252 Paris Cedex 05 (France)

yining.hu@imj-prg.fr

Abstract

We give an explicit implicit function theorem for formal power series that is valid for all fields, which implies in particular Lagrange inversion formula and Flajolet-Soria coefficient extraction formula known for fields of characteristic 0.

Theorem 1. *Let K be an arbitrary field. If $P(X, Y) \in K[[X, Y]]$ and $f(X) \in K[[X]]$ are such that $f(0) = 0$, $P(X, f(X)) = f(X)$ and $P'_Y(0, 0) = 0$. Then*

$$[X^n]f = \sum_{m \geq 1} [X^n Y^{m-1}] (1 - P'_Y(X, Y)) P^m(X, Y).$$

If the characteristic of K is 0, we also have the following form

$$[X^n]f = \sum_{m \geq 1} \frac{1}{m} [X^n Y^{m-1}] P^m(X, Y).$$

Remark 1. The conditions $P(X, f(X)) = 0$ and $f(0) = 0$ imply that $P(0, 0) = 0$. As $P'_Y(0, 0)$ is also 0, the sums in both expressions of $[X^n]f$ are finite.

Remark 2. When $P(X, Y) = X\phi(Y)$, where $\phi(X) \in K[[X]]$ and $\phi(0) \neq 0$, we obtain the Lagrange inversion formula

$$[X^n]f = [Y^{n-1}] (\phi(X)^n - Y\phi'(X)\phi(X)^{n-1}).$$

If the characteristic of K is 0, we also have the following form

$$[X^n]f = \frac{1}{n} [Y^{n-1}] \phi(Y)^n.$$

Remark 3. When $P(X, Y)$ is a polynomial in X and Y , we obtain a generalisation of Flajolet-Soria coefficient extraction formula [2].

Definition 1. Let $P(X, Y) \in K[[X, Y]]$,

$$P(X, Y) = \sum_{j=0}^{\infty} a_j(x) Y^j,$$

where $a_j(x) \in K[[X]]$. We define

$$P^{[m]}(X, Y) = \sum_{j=m}^{\infty} \binom{j}{m} a_j(x) Y^{j-m}.$$

for $m \in \mathbb{N}$.

The motivation of the definition of $P^{[m]}$ is to avoid the factorials in the denominators in the Taylor series, which do not make sense in positive characteristic. Once this obstacle is circumvented, the Taylor formula works as expected.

Proposition 1. *Let K be an arbitrary field. Let $P(X, Y) \in K[[X, Y]]$ and $f(X) \in K[[X]]$ with $f(0) = 0$. Then*

$$P(X, Y) = \sum_{m=0}^{\infty} (Y - f(X))^m P^{[m]}(X, f(x)). \quad (*)$$

Proof. Let $P(X, Y) = \sum_{j=0}^{\infty} a_j(X) Y^j$ with $a_j(X) \in K[[X]]$ for $j \in \mathbb{N}$. We prove that for all $k \in \mathbb{N}$, the coefficient of Y^k in the left side and right side of $(*)$ is equal. Indeed, we have

$$\begin{aligned} & [Y^k] \sum_{m=0}^{\infty} (Y - f(X))^m P^{[m]}(X, f(x)) \\ &= \sum_{m=k}^{\infty} \binom{m}{k} (-f(X))^{m-k} \sum_{j=m}^{\infty} \binom{j}{m} a_j(x) f(X)^{j-m} \\ &= \sum_{j=k}^{\infty} a_j(X) f(X)^{j-k} \sum_{m=k}^j \binom{j}{m} \binom{m}{k} (-1)^{m-k} \\ &= \sum_{j=k}^{\infty} a_j(X) f(X)^{j-k} \sum_{m=k}^j \binom{j}{j-m, m-k, k} (-1)^{m-k} \\ &= a_k(X). \end{aligned}$$

We have the last equality because $\sum_{m=k}^j \binom{j}{j-m, m-k, k} (-1)^{m-k} = 1$ if $j = k$ and 0 if $j > k$. This is because we have the multinomial expansion

$$\begin{aligned} (a + b + c)^j &= \sum_{k \leq m \leq j} \binom{j}{j-m, m-k, k} a^{j-m} b^{m-k} c^k \\ &= \sum_{k=0}^j \sum_{m=k}^j \binom{j}{j-m, m-k, k} a^{j-m} b^{m-k} c^k \end{aligned}$$

When we take $a = 1$, $b = -1$, the identity becomes

$$c^j = \sum_{k=0}^j \left(\sum_{m=k}^j \binom{j}{j-m, m-k, k} (-1)^{m-k} \right) c^k.$$

Seeing this as a polynomial identity in the variable c gives us the desired result. \square

The following corollary is immediate.

Corollary 1. *Let K be an arbitrary field. Let $Q(X, Y) \in K[[X, Y]]$ and $f(X) \in K[[X]]$ be such that $f(0) = 0$ and $Q(X, f(X)) = 0$. Then there exists $R(X, Y) \in K[[X, Y]]$ such that $Q(X, Y) = (Y - f(X))R(X, Y)$.*

Definition 2. For the formal power series in $K((X_1, \dots, X_m))$

$$P(X_1, X_2, \dots, X_m) = \sum_{n_i > -\mu} a_{n_1 n_2 \dots n_m} X_1^{n_1} X_2^{n_2} \dots X_m^{n_m}$$

its (principal) diagonal $\mathcal{D}f(t)$ is defined as the element in $\kappa((T))$

$$\mathcal{D}P(T) = \sum a_{nn \dots n} T^n.$$

The following proposition is a generalization of Proposition 2 from [1], the only difference in the proof is in the first step where we use Corollary 1 to factorize $Q(X, Y)$.

Proposition 2. *Let K be an arbitrary field. Let $Q(X, Y) \in K[[X, Y]]$ and $f(X) \in K[[X]]$ be such that $f(0) = 0$, $Q(X, f(X)) = 0$ and $Q'_Y(0, 0) \neq 0$, then*

$$f(X) = \mathcal{D} \left(Y^2 \frac{Q'_Y(XY, Y)}{Q(XY, Y)} \right).$$

Proof. Using Corollary 1 we can write $Q(X, Y) = (Y - f(X))R(X, Y)$ with $R(X, Y) \in K[[X, Y]]$. We have $R(0, 0) \neq 0$ because $f(0) = 0$ and $Q'_Y \neq 0$. Then

$$\frac{1}{Q(X, Y)} Q'_Y(X, Y) = \frac{1}{Y - f(X)} + \frac{R'_Y(X, Y)}{R(X, Y)}.$$

Replacing X by XY and multiplying by Y^2 we get

$$\mathcal{D} \left(Y^2 \frac{Q'_Y(XY, Y)}{Q(XY, Y)} \right) = \mathcal{D} \left(\frac{Y^2}{Y - f(XY)} \right) + \mathcal{D} \left(Y^2 \frac{R'_Y(XY, Y)}{R(XY, Y)} \right). \quad (\dagger)$$

For the first term on the right side of (\dagger) we have

$$\begin{aligned} & \mathcal{D} \left(\frac{Y^2}{Y - f(XY)} \right) \\ &= \mathcal{D} \left(\frac{Y}{1 - Y^{-1}f(XY)} \right) \\ &= \mathcal{D} \left(\sum_{n=0}^{\infty} Y^{-n+1} f(XY)^n \right) \\ &= \mathcal{D} (f(XY)) \\ &= f(X). \end{aligned}$$

For the second term, as $R(0, 0) \neq 0$, $\frac{R'_Y(XY, Y)}{R(XY, Y)}$ is a power series in XY and Y , so when we multiply this by Y^2 there is no diagonal term. □

Proof of Theorem 1. Let the power series $Q(X, Y)$ be defined as $Q(X, Y) = P(X, Y) - Y$, then $Q'_Y(0, 0) = P'_Y(0, 0) - 1 \neq 0$, and $Q(X, f(X)) = P(X, f(X)) - f(X) = 0$. According to Proposition 2,

$$\begin{aligned}
f &= \mathcal{D}\{Y^2 Q'_Y(XY, Y)/Q(XY, Y)\} \\
&= \mathcal{D}\{Y^2(P'_Y(XY, Y) - 1)/(P(XY, Y) - Y)\} \\
&= \mathcal{D}\{Y(1 - P'_Y(XY, Y))/(1 - \frac{P(XY, Y)}{Y})\} \\
&= \mathcal{D}\{Y(1 - P'_Y(XY, Y))(1 + \sum_{m \geq 1} (\frac{P(XY, Y)}{Y})^m)\}.
\end{aligned}$$

We have the last equality due to the fact that $P'_Y(0, 0) = 0$, $\frac{P(XY, Y)}{Y}$ has no constant term, and therefore $1/(1 - \frac{P(XY, Y)}{Y}) = 1 + \sum_{m \geq 1} (\frac{P(XY, Y)}{Y})^m$. As in each term of $Y(1 - P'_Y(XY, Y))$ the power of Y is larger than that of X , it cannot contribute to the diagonal. Therefore,

$$\begin{aligned}
f_n &= [X^n Y^n] Y(1 - P'_Y(XY, Y))(1 + \sum_{m \geq 1} (\frac{P(XY, Y)}{Y})^m) \\
&= [X^n Y^n] Y(1 - P'_Y(XY, Y))(\sum_{m \geq 1} (\frac{P(XY, Y)}{Y})^m) \\
&= \sum_{m \geq 1} [X^n Y^{m-1}] (1 - P'_Y(X, Y)) P(X, Y)^m.
\end{aligned}$$

□

References

- [1] H. Furstenberg, “Algebraic functions over finite fields”, J. Algebra 7, 271–277 (1967).
- [2] M. Soria, Thèse d’habilitation (1990), LRI, Orsay.